PB90-100900

Membership Values as Indicators of Complications in Chromatoraphy

(U.S.) Environmental Protection Agency, Research Triangle Park, NC

1989

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before comple

1. REPORT NO.	2.	3.
EPA/600/J-89/087		PB90-100900

4. TITLE AND SUBTITLE
Membership Values as Indicators of Complications in Chromatography

5. REPORT DATE

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Dr. Joseph F. Walling

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
U. S. Environmental Protection Agency
Atmospheric Research and Exposure Assessment Laboratory
Methods Research and Development Division
Research Triangle Park, NC 27711

10. PROGRAM ELEMENT NO.
ANMA1E

11. CONTRACT/GRANT NO.

12. SPONSORING AGENCY NAME AND ADDRESS
U. S. Environmental Protection Agency
Office of Research and Development
401 M Street, S.W.
Washington, DC 20460

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE
EPA/600/09

15. SUPPLEMENTARY NOTES
Journal Article --- To be submitted to Journal of Chromatography

16. ABSTRACT

Fuzzy set theory is used to provide a figure of merit, the membership value, for data sets as an alternative to censoring databases. Applications to distributed air volume sets and pairs of suitable chromatograms are illustrated.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	8
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

Journal of Chromatography, 473 (1989) 267–272
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

EPA/600/J-89/087

CHROM. 21 550

Note

Membership values as indicators of complications in chromatography

J. F. WALLING

Atmospheric Research and Exposure Assessment Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711 (U.S.A.)

(First received December 6th, 1988; revised manuscript received April 6th, 1989)

Chromatograms from environmental samples, including those from ambient air, are usually complex. The common presence of coeluting species in such chromatograms and the interpretive problems they pose even when using mass spectral detection with non-interactive computer based data processing are recognized and have been discussed[1–3]. Additional factors which can frustrate the straightforward use of analytical results from adsorbent sampling have also been presented[4].

In such situations data sets in which parameters are distributed can be useful; *e.g.*, chromatograms of one sample from two columns which produce poorly correlated retention times, or multiple adsorbed samples collected simultaneously but with a distribution of air volumes[5]. Practical limits on the number of multiple analyses are understandably severe especially in large-scale applications; but even if such data are available, their usual handling based on statistics can result in the reporting of severely censored databases. An alternative to that, to totally uncensored databases or to those with mixed or unspecified censoring is desirable, especially when those who produce and those who use the data are not the same.

Assuming competent and competently performed procedures, the difficulty becomes one of assigning any datum to a class of essentially uncomplicated results or to others containing data affected by complicating processes. Assignments based on identification of any specific complication with or without a correction for its effect on the observations are generally impractical. Statistically based assignments are not only weak when sets contain few samples, probably only two, but they are binary (membership values in a class can be only zero or one). Severe censoring can result. If multiple grades of membership to the class of uncomplicated results could be obtained from the data themselves, those grades might be part of a report of the uncensored data and make an explicit statement about the conformance of each data pair to expectations for complication free results.

Fuzzy set theory was suggested as an approach for situations such as this where uncertainty about the criteria for class membership exists[6]. Although published chemical applications are not presently numerous, a relevant summary of the fuzzy perspective with some chemical applications can be found elsewhere[7]. Here in this very simple application, a fuzzy relationship is defined to express the "nearness" of pairs of relevant measurements by a membership value in the interval zero to one. A value of one indicates equality and implies a complication free result. Decreasing values imply

© 1989 Elsevier Science Publishers B.V.

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

increasing inequality and, therefore, increasing effects of some complications. Two illustrations of this approach follow; one using sets of atmospheric samples adsorbed on Tenax and the other using a synthetic mixture analyzed on serially coupled columns under two different sets of conditions.

EXPERIMENTAL[a]

Data

The data from sets of atmospheric samples adsorbed on Tenax and analyzed by gas chromatography–mass spectrometry have been reported previously (see Table 1 in ref. 5). Only data from the front tubes are used here.

Synthetic gaseous mixtures for the chromatographic experiments were prepared by neat liquid injection and subsequent volatilization into a suitable vessel[8]. Expected retention times and flame ionization detector responses were determined for a calibration mixture of hydrocarbons and halogenated hydrocarbons of some environmental interest. The test mixture of 75 species contained in addition to those, a hydrocarbon mixture not characterized but containing gasoline components.

Separations were done using a Perkin-Elmer Model 8500 gas chromatograph with serially coupled 30 m × 0.32 mm I.D. fused-silica columns; 0.2-μm DB-1 followed by 0.2-μm DB-210. Control of helium pressures at the inlet and column coupling points determined relative residence times of the carrier in each column. Two sets of pressure settings were chosen arbitrarily. Calibration and test mixtures were analyzed at each setting. In all cases temperatures were programmed at 4°C/min starting at 40°C at the time of injection.

Peak areas and locations were determined by the chromatograph's data system and identified with a substance in a conventional way. The ratios of detector outputs from the test mixture to those from calibrations are shown in Tables II and III as normalized areas.

Fuzzy approach

Expectations were modeled as fuzzy relations. Assuming no complications, all Tenax samples in a set should yield nearly equal concentrations for any given substance. Similarly, in these chromatographic experiments, each substance should appear near its expected retention times and the quantities found should be nearly equal as well.

Membership values (MV) for pairs of observations in the Tenax sets were calculated using a one-dimensional Gaussian function:

$$\mathrm{MV}_i = \exp - [(x_{ij} - x_{ik})^2/2s_i^2]$$

where x_{ij} and x_{ik} are the pairs of concentrations for substances i and s_i is a scale factor. Chromatographic data were described by a membership function of the same form but in three or in some cases four dimensions; two or three retention time differences and one difference in areas. Computations are straightforward manipulations of data as shown in the argument of the one-dimensional function or a sum of such terms to

[a] Mention of trade names does not constitute endorsement or recommendation for use.

comprise the argument of the multidimensional function. They were done here on a Compaq Deskpro 286 using PC-Matlab; however, any spreadsheet or other program capable of such manipulations would suffice.

RESULTS AND DISCUSSION

Substantial latitude exists in the choice of the membership function and its parameters' values. Different choices will yield somewhat different membership values. Choice of the Gaussian was for convenience both in computations and selecting the scale factors, s_i. Those used were 0.15 min for retention time terms and

TABLE I

MEMBERSHIP VALUES (MV) AND NET APPARENT ATMOSPHERIC CONCENTRATIONS ($\mu g/m^3$) FROM THREE DISTRIBUTED AIR VOLUME SETS

	MV Mx Mn	*MV 6 & 27*	*Air volume sampled (l)*			
			6	*11.4*	*26.9*	*54*
Benzene	0.00	0.00	7.0	4.9	3.9	2.1
Methylbenzene	0.00	0.92	57.0	50.4	53.6	20.7
1,2-Dimethylbenzene	0.00	0.01	3.8	3.8	2.4	1.9
Ethylbenzene	0.00	0.00	3.4	2.7	2.0	1.6
Trichloroethene	0.00	0.00	0.65	0.44	0.39	0.22
1,1,1-Trichloroethane	0.00	0.00	5.5	2.9	2.1	1.0
Tetrachloroethene	0.11	0.95	2.2	2.0	2.1	1.6
Benzaldehyde	0.00	0.00	32.4	20.0	11.1	5.6
1-Phenylethanone	0.00	0.00	32.4	29.3	10.4	9.1
Benzonitrile	0.00	0.00	2.9	1.7	1.0	0.0
Benzene	0.27	0.83	17.1	15.9	15.6	13.4
Methylbenzene	0.00	0.00	122	126	22.6	10.4
1,2-Dimethylbenzene	0.63	0.72	9.6	9.3	8.5	8.3
Ethylbenzene	0.98	0.99	6.3	6.1	6.2	6.1
Ethenylbenzene	0.28	0.85	1.2	1.2	1.1	1.4
Trichloroethene	0.71	0.98	3.4	3.3	3.3	3.0
1,1,1-Trichloroethane	0.00	0.18	8.6	8.8	6.5	4.2
Tetrachloroethene	0.82	0.82	38.4	35.8	34.9	35.1
1,2-Dichloropropane	0.70	0.77	18.5	19.2	20.6	21.0
Benzaldehyde	0.00	0.00	11.3	10.7	5.6	4.6
1-Phenylethanone	0.00	0.00	0.39	0.14	0.00	3.2
Benzene	0.00	0.01	4.6	3.7	2.8	2.5
Methylbenzene	0.08	0.89	7.9	9.4	8.5	6.7
1,2-Dimethylbenzene	0.00	0.01	2.6	2.1	1.6	1.3
Ethylbenzene	0.41	1.00	1.0	1.1	1.0	0.90
Ethenylbenzene	0.00	0.00	0.75	0.85	0.41	0.23
Trichloroethene	0.00	0.03	0.39	0.37	0.26	0.17
1,1,1-Trichloroethane	0.00	0.92	1.7	1.9	1.6	0.80
Tetrachloroethene	0.77	0.97	2.8	2.9	2.7	2.6
1,2-Dichloropropane	0.27	0.36	2.9	3.2	3.6	3.7
Benzaldehyde	0.00	0.00	27.7	16.3	6.1	4.0
1-Phenylethanone	0.00	0.00	34.4	0.00	6.8	6.0
Benzonitrile	0.00	0.00	2.3	1.4	0.00	0.00

15% of the average concentration (Table I) or average normalized area (Tables II and III) taken as reasonable standard deviations of these differences. Although the considerations which shaped these choices are chemically plausible and statistically based, the membership values are not statistical results.

If distributed air volume samples are taken at all, sets of two are more likely to be collected than the sets of four shown in Table I. Therefore, a treatment using only two of the four samples was done. Two different membership values are shown; one using the extreme values for each substance at whatever sampling volume they appear (perhaps the easiest way to assign the whole set of four to a class) and the other using just the results from the 6- and 26.9-l samples as a more usual application.

Although the values reflect the behavior of the relevant observations, inspection will reveal some striking differences as well as similarities resulting from these choices. This underscores that each membership value describes only the fuzzy relationship of the data pair. Clearly, any indicator of the importance of complications discernable from the data (here, only factors with a strong non-linear dependence on air volume) must depend critically on the data used.

TABLE II

MEMBERSHIP VALUES (MV) AND OBSERVATIONS FOR 24 ANALYTES FROM PAIRS OF CHROMATOGRAMS OF A 75-COMPONENT MIXTURE

	MV	*Chromatogram A*			*Chromatogram B*		
		Expected retention time (min)	*Observed retention time (min)*	*Normalized area*	*Expected retention time (min)*	*Observed retention time (min)*	*Normalized area*
2-Bromo-1-propene	0.00	3.89	3.92	2.04	6.00	6.03	0.90
1-Bromopropane	0.00	4.89	4.88	1.94	7.44	7.48	1.15
Bromodichloromethane	0.00	6.35	6.40	4.83	9.66	9.76	8.26
2,2-Dichlorobutane	0.00	7.52	7.49	1.05	11.00	11.04	2.54
Methylbenzene-d_8	0.80	8.51	8.48	1.03	12.34	12.40	1.11
2,2-Dibromopropane	0.72	8.70	8.67	1.04	12.56	12.64	1.13
1,3-Dichlorobutane	0.78	11.12	11.10	1.00	14.98	15.05	1.08
1,1,2-Trichloropropane	0.82	11.29	11.27	1.00	15.52	15.59	0.94
1,1,1,2-Tetrachloroethane	0.69	11.40	11.37	0.98	15.89	15.96	1.08
1,4-Dichlorobutane	0.00	14.31	14.28	0.99	18.30	18.36	4.74
1,2,3-Trichloropropane	0.00	14.45	14.42	0.97	18.82	18.88	2.71
(1-Methylethyl)benzene	0.86	14.70	14.68	0.99	19.75	19.83	1.00
1,1,1,2-Tetrachloropropane	0.87	15.06	15.04	1.00	20.06	20.13	0.97
1,3-Dibromopropane	0.83	15.58	15.54	1.01	20.20	20.27	0.97
1-Methyl-2-chlorobenzene	0.81	15.98	15.95	1.01	21.05	21.12	1.08
1-Methyl-4-chlorobenzene	0.87	16.31	16.28	1.02	21.31	21.38	0.99
1,3-Dibromobutane	0.00	17.59	17.49	3.38	22.43	22.50	0.96
1,4-Dichlorobenzene	0.69	18.52	18.49	0.98	23.76	23.84	0.89
Benzonitrile	0.73	19.28	19.23	0.98	22.61	22.67	0.90
1,2-Dichlorobenzene	0.79	19.62	19.60	0.96	24.90	24.97	0.89
Butylbenzene	0.16	20.37	20.36	1.01	25.61	25.68	0.76
1,4-Dibromobutane	0.35	21.20	21.18	1.07	26.18	26.25	0.87
1-Ethenyl-4-chlorobenzene	0.72	21.28	21.28	0.92	26.66	26.73	0.83
1-Phenylethanone	0.60	22.30	22.26	0.83	26.40	26.46	0.73

TABLE III

MEMBERSHIP VALUES (MV) AND OBSERVATIONS FOR ANALYTES COELUTING IN ONE OF A PAIR OF CHROMATOGRAMS

		Chromatogram A			*Chromatogram B*		
		Expected retention time (min)	*Observed retention time (min)*	*Normalized area*	*Expected retention time (min)*	*Observed retention time (min)*	*Normalized area*
1-Bromo-2-chloroethane	0.01	7.17	7.12	3.40	10.45	10.55	1.76
1-Chloro-2,3-epoxypropane		7.92	7.88	1.07			
Dibromochloromethane	0.90	8.96	8.94	1.06	13.17	13.23	1.13
3,4-Dichloro-1-butene		9.38	9.35	1.20			
Tribromomethane	0.83	12.29	12.26	1.00	17.18	17.26	1.00
1,4-Dimethylbenzene		12.48	12.45	1.02			
1,1,2,2-Tetrachloroethane	0.00	13.64	13.61	3.39	18.08	18.15	0.98
1,4-Dichloro-2-butene		13.87	13.84	1.05			
1,3,5-Trimethylbenzene	0.87	16.62	16.60	1.05	21.84	21.91	0.99
1,2,2,3-Tetrachloropropane		16.75	16.72	0.92			

Illustrative data for some of the analytes in the two chromatograms of the gaseous test mixture are shown in Tables II and III. Table II contains results for those analytes which, during calibration, eluted singly under both sets of conditions. These are describable by the three-dimensional membership function. Normalized areas differing greatly from one indicate coeluents from the hydrocarbons and are reflected appropriately in the membership values. Even though mixtures were chosen to cause clear effects from coelution, a membership value range is exhibited. Even more subtle gradations are to be expected from environmental samples with a greater range of causes.

Optimization of the two sets of conditions was deliberately avoided to insure that some potential analytes coeluted. Table III displays some that separated under just one set of conditions and can be described by a membership function of four dimensions. (Because of the way the normalization was done here, in a complication free case, the sum of individual areas should equal twice the area when they coeluted.) In these instances, all of the results from the separated species could be reported accompanied by the membership value for the pair. These values reflect conformance to the expectation that all observations are free of further complications. Clearly, the same idea can be applied if larger numbers coelute under one set of conditions.

CONCLUSIONS

Membership values calculated quite simply using a fuzzy approach can reflect the relative importance of some common complications in analyses of complex mixtures. The membership value depends on the choice of membership function, values of its parameters and the data pairs. In this application, the function and its parameters' values must be made explicit. Any inadequacies in the pairs of experiments which fail to test the relationships (*e.g.*, nearly equal air volumes or columns too similar in separation characteristics) would not be apparent; however, in an

application to chromatogram pairs, the adequacy of the retention time observations (sometimes overlooked by data users) are included. Within these limits, each value can be a useful index of complications eliminating the need for the sometimes severe censoring of data.

ACKNOWLEDGEMENTS

The author wishes to thank L. H. Wright and J. V. Daughtridge for their generosity and excellent work in obtaining the chromatograms. This article has not been subjected to Agency review and, therefore, does not necessarily reflect the views of the Agency. No official endorsement should be inferred.

REFERENCES

1 D. Rosenthal, *Anal. Chem.*, 54 (1982) 63.
2 J. M. Davis and J. C. Giddings, *Anal. Chem.*, 57 (1985) 2168 and 2178.
3 M. Martin, D. P. Herman and G. Guiochon, *Anal. Chem.*, 58 (1986) 2200.
4 J. F. Walling, *Atmos. Environ.*, 18 (1984) 855.
5 J. F. Walling, J. E. Bumgarner, D. J. Driscoll, C. M. Morris, A. E. Riley and L. H. Wright, *Atmos. Environ.*, 20 (1986) 51.
6 L. A. Zadeh, *Inform. Control*, 8 (1965) 338.
7 H. Bandemer and M. Otto, *Mikrochim. Acta*, 2 (1987) 93.
8 C. Morris, R. Berkley and J. Bumgarner, *Anal. Lett.*, A20 (1983) 1585.